iJRASET
International Journal For Research in
Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 14 Issue: VI Month of publication: June 2026

DOI: https://doi.org/10.22214/ijraset.2026.83615

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue VI June 2026- Available at www.ijraset.com

AI-Driven Real-Time Emotion Detection Using an Explainable CNN-GCN Hybrid Model

Priyanka Kudithetti[1], B. Kranthi Kiran[2]
[1]*M. Tech Student Department of Computer Science and Engineering, JNTU Hyderabad, Telangana*
[2]*Professor of Computer Science and Engineering, JNTU Hyderabad, Telangana*

Abstract: Facial emotion recognition is an important component of intelligent systems and has applications in areas such as human–computer interaction, healthcare, surveillance, and behavioral analysis. Although deep learning techniques have achieved significant progress in this field, many existing approaches mainly rely on facial appearance features and often overlook the structural relationships between facial landmarks, which can affect their performance in real-time environments. To address this challenge, this study presents an AI-driven real-time emotion recognition framework based on a hybrid Convolutional Neural Network (CNN) and Graph Convolutional Network (GCN). The proposed model combines a pretrained ResNet18 network for extracting facial appearance features with a GCN module that captures geometric relationships from 468 facial landmarks obtained using MediaPipe Face Mesh. To improve feature learning and model robustness, geometric loss and adversarial feature regularization are incorporated into the training process. The framework was trained and evaluated on the CK+ dataset containing seven emotion categories and compared with several state-of-the-art models, including VGG16, VGG19, ResNet18, MobileNetV3, ConvNeXt, and EfficientNet-B0. Experimental results show that the proposed CNN-GCN model achieved an accuracy of 99.2%, while ResNet18 and EfficientNet-B0 achieved the highest accuracy of 99.6%. To support practical deployment, the system was implemented as a Flask-based web application capable of real-time webcam and image-based emotion recognition. In addition, explainable AI techniques, including Grad-CAM and LIME, were integrated to provide visual insights into model predictions. The proposed framework offers an accurate, interpretable, and reliable solution for real-time emotion-aware intelligent applications.
Keywords: Facial Emotion Recognition, Convolutional Neural Network (CNN), Graph Convolutional Network (GCN), Explainable Artificial Intelligence (XAI), Real-Time Emotion Detection, Human–Computer Interaction.

I. INTRODUCTION

Facial emotion recognition is a research field of artificial intelligence, computer vision, and human-computer interaction because of the system's ability to understand human emotional behavior. Automated emotion by observation of facial expressions. Online health education smart behavior driving virtual. Improvements in face emotion detection using deep learning methods by representing and classifying the features [1]. Fusion learning methods have greatly improved the adaptability and robustness of emotion detection in real-world complex environments [2]. Face emotion recognition has increased the importance of attention-based feature extraction networks [3]. In recent years, a large number of deep learning-based facial expression detection studies have highlighted the importance of developing reliable emotion detection techniques [4]. Recently, many facial emotion recognition systems have been proposed and achieved impressive performances in controlled environments, however, they still contain a number of shortcomings, which can affect deployment challenges, prediction accuracy, etc. Most existing approaches are based on appearance features and they fail to model the inter-point geometry relationship among landmark points in the face, leading to poor performance under the challenges of pose variation, light variation, occlusion, and intensity variations. The graph-based relationship learning approaches have been proposed for modeling the spatial dependency among face regions [5]. Similarly, the geometry-guided facial representations approaches have been proposed for pose-invariant expression recognition [6]. However, most of the real-time emotion recognition frameworks suffer from unstable predictions in different contexts, and lack adaptability to challenging environments [7]. Further, most of the existing approaches lack interpretability and are vulnerable to adversarial attacks [8]--[10].

Some intellectual challenges are discussed in real-time facial expression recognition, including emotion prediction reliability, interpretability, and effectiveness of real-time deployment. Then, an intelligent real-time facial emotion recognition framework is presented to improve the expression recognition accuracy, robustness, interpretability, and real-time deployment effectiveness. The developed framework is a suitable tool for reliable visual appearance facial expression recognition and structural facial relationship expression recognition in real-time real-time monitoring scenarios.

Enhancing the estimation interpretability and real-time real-time deployment efficiency of the developed framework has a good effect on the design of intelligent emotion-aware behavior analysis state systems, reliable healthcare support state systems, intelligent transportation systems, smart learning environments, intelligent security systems, and the creation of emotion-aware performance user interfaces. These also benefit the intelligence system design and development of high state facial expression recognition. This is because the developed system adopts multiple learning state recognition strategies to improve the recognition accuracy of emotional state expressions. Additionally, it leverages multiple learning state interpretation to enhance the consistency of state feature representation, ensuring the reliability of state expression recognition. The framework is implemented, and experimental results are used to demonstrate the effectiveness of the designed system.

II. RELATED WORK

Deep learning has become a powerful tool to recognize human facial emotions, with recent significant progress in this area. Burrows et al. proposed a real-time emotional reflective user interface that utilized deep convolutional networks and generative adversarial learning for interactive emotional analysis [11]. Janjanam et al. introduced a real-time emotion analytics system for monitoring multiple people in real time and visualizing the prediction accuracy [12]. Sadiq et al. proposed a real-time face emotion recognition system that is able to deal with occlusion using an occlusion-adaptive emotion detection framework [13]. While various methods have been proposed for real-time facial emotion recognition, it is critical to design a deep learning framework that can incorporate with real-time interpretability, consistency in prediction and human emotional understanding of facial emotion images so as to overcome the remaining challenges in this domain.

The facial emotion recognition task requires the ability to capture and interpret subtle variations in a person's inner feelings as reflected on their facial expression. In general, such variations manifest in the face in a form of very subtle changes in facial landmarks location.

A number of techniques have been proposed for capturing the facial muscle movements producing facial expressions, including variations in mouth shape, eye shape, eyebrows, and the head pose. Owing to heterogeneity of facial expression across different populations, ethnicity, and gender, developments in this area has involved various aspects, including representation of facial features, temporal modeling of facial expressions, the development of emotion sets, and emotion classification techniques. In addition, the role of studies in emotion expression has been explored, as well, in capturing subtler variations in facial expressions, as well as providing realizable systems for detecting emotion from expression. This paper provides a review on analytic modeling in facial emotion recognition, with emphasis on recent developments in facial feature representation and expression dynamics. Recent research has also investigated advanced facial representation and spatial understanding techniques to enhance emotional analysis performance. Ropero et al. investigated emotion recognition using three-dimensional point cloud representations generated from facial expression images [14]. Their framework demonstrated improved structural facial understanding, although higher computational requirements limited lightweight real-time deployment.

Human perceptions of intelligence increasingly rely on transparency and interpretability. Selvaraju et al. proposed Grad-CAM as a visual explanation method by drawing relevant image segments involved in the prediction [17]. Ribeiro et al. proposed the LIME framework, which provided a way to generate local explanations of a classifier [18]. Both methods greatly promoted the explainable artificial intelligence field. But most facial emotion recognition systems are still blind about the interpretability, which is integrated into the real-time system, where the workloads of interpretability can be reduced and the risk of misbehavior can be low.

Facial landmark extraction remains to be one of the vital factors for the stability of facial landmark detection and dataset representation. Lugaresi et al. proposed thereby a MediaPipe perception framework for accurate facial landmark extraction and real-time processing applications [19]. The framework proposed in this way has a good facial representation functionality for interactive environment. Lucey et al. propose a CK+ dataset with emotion-specific facial expressions and action unit annotations and it is one of the most used datasets for benchmarking facial emotion recognition evaluations [20]. However, there are still some challenges left, such as scarcity of geometric facial representation integration, lack of interpretability, instability of real-time prediction, and lack of robustness.

To observe these limitations, the proposed framework focuses on the comprehensive vision and structure of facial features to improve the reliability, interpretability, and stable operation of real-time emotion recognition. Besides, it also focuses on the transparent understanding of predictions, better handling of practical expression facial variations, thus making a meaningful contribution towards building more reliable and interactive emotion-aware intelligent systems for realistic scenarios.

III. METHODOLOGY

A. System Overview

The proposed framework was designed to perform real-time facial emotion recognition by combining deep learning and graph-based learning techniques. Using the CK+ dataset with seven emotion categories, the system extracts facial appearance features through a pretrained ResNet-based CNN and captures geometric relationships among 468 facial landmarks using a Graph Convolutional Network (GCN). To improve classification accuracy and feature learning, a hybrid loss function incorporating cross-entropy and geometric loss was employed. The framework was further deployed as a Flask-based web application supporting real-time webcam and image-based emotion recognition, while Grad-CAM and LIME were integrated to provide explainable and interpretable predictions. Experimental results demonstrated the framework's ability to deliver accurate, reliable, and real-time emotion recognition.

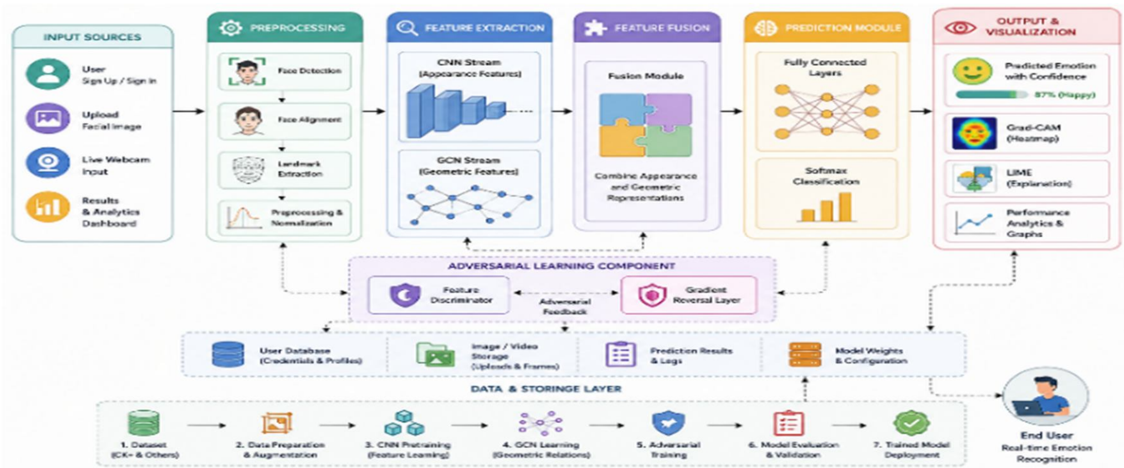


Fig.1 System Architecture

Fig. 1 shows the theoretical process of the proposed real-time facial emotion recognition framework, which contains the real-time deployment framework, the framework of the feature learning and fusion module with CNN and GCN, the classification module, and the visualization module with Grad-CAM and LIME of explainable AI. The proposed framework analyzes the facial appearance and geometry representations of the face jointly. It can be deployed with Flask-based and run in real times, and provide a flexible platform for intelligent emotion applications with web-based applications through a webcam-based real-time deployment environment and images for emotion predictions. And Grad-CAM and LIME are used to recognize the highly salient facial regions that cause the model to classify as the target emotion.

B. Data Acquisition

The facial emotion recognition framework has been implemented and tested on the FER2013 image dataset that is publicly available in the Kaggle repository. The FER2013 dataset consists of 1,144 images that have been labeled including seven facial expressions: anger, disgust, fear, happiness, sadness, surprise, and neutral. The FER2013 dataset includes 245 images in the validation set and 246 images in the testing set. The FER2013 dataset is used for facial expression recognition because the dataset includes faces with variations in illumination conditions, expression intensity, and facial orientations. Balanced training, validation, and testing set can ensure that the emotion recognition framework will generalize well for the real-time usage.

C. Preprocessing

Preprocessing processes were applied to increase data performance, augment learning consistency, reduce class imbalances and improve facial representations, thereby facilitating accurate, reliable and robust real-time facial emotion recognition performance.

1) *Dataset Preparation and Organization:* The Cohn-Kanade (CK+) Extended Dataset was categorized into seven different emotional classes namely anger, contempt, disgust, fear, happiness, sadness, and surprise. The image path and class information were organized into data collections so that the dataset could be prepared for subsequent training, validation, and testing procedures. The categorization of the dataset made it easy for data handling and improved the consistency between training and testing sets. The organized Cohn-Kanade dataset also provided reliable performance of the learning process.

2) *Train–Validation–Test Data Partitioning:* The complete dataset was separated into training, validation and testing subsets to improve the model generalization and prevent overfitting. Stratified partitioning was performed accordingly. The training, validation and testing sub-datasets contained 1,144, 245, and 246 samples, respectively. Appropriate data set partitioning further improves the reliability of analysis results, prevention of overfitting, and success of testing the emotion recognition ability on realistic facial expressions.
3) *Class Imbalance Mitigation:* During model training, the imbalanced distribution among emotion classes was countered with an adaptive weighted sampling approach so that minority emotion classes receive a proportionally higher weighting. During the training process, an adaptive weighting strategy was applied so that minority emotion classes receive a higher weighting, based on the imbalanced distribution among emotion classes. The balancing strategy particularly contributed to the improvement of the classification performance on the under-represented classes, such as contempt and fear, because it worked well with the weighted random sampling method to improve the sample distribution. Our experiments verified that the weighting and sampling strategies for counterbalancing distribution among emotion categories is meaningful, and improved our classification system in achieving a higher overall recognition rate. The weighted random sampling approach, especially to counter balance the distribution among emotion categories, promotes a more fair training and improves the overall recognition accuracy.
4) *Image Preprocessing and Spatial Transformation:* Facial images were utilized in a feature learning task for preprocessing purposes including scaling, normalizing and converting them to tensors in order to acquire a consistent input representation. Image normalization allowed numerical consistency during feature learning while increasing image resolution normalized all training samples to a specific spatial resolution. Uniform preprocessing allowed for consistent image representation which improved convergence during training and allowed for the efficient extraction of facial expression characteristics for accurate and reliable emotion recognition.
5) *Data Augmentation:* To increase the diversity of the dataset and resilience of learning, we employed some data augmentations, such as rotational transformation, random horizontal flip and color variation correction. These variations mimic realistic face appearance changes, which will reduce the sensitivity of model to changes like different orientations, illumination, and expression intensities. When a large number of novel facial situations are encountered, the augmented samples of training data help the framework to generalize better which also lower the overfitting caused by the lack of training diversity. With augmented data, the stability of emotion recognition and real-time prediction in dynamic environment were considerably aided.
6) *Facial Landmark Extraction and Geometric Feature Engineering:* By extracting 468 two-dimensional facial landmark coordinates with MediaPipe Face Mesh, a geometric facial representation was developed. Landmark localization was used to acquire the precise structural face characteristics needed for geometrical emotion analysis like the eyes, eyebrows, nose, lips, and jawline. Converting the extracted coordinates into structured numeric representations facilitated learning of spatial face relationships and subtle change in expression. Geometrical feature engineering contributed to enhancing the framework's ability to comprehend structural face dependencies beyond appearance-based data. Accurate landmark extraction also helped improved robustness under various facial orientations and intensity of emotional expression.

D. Algorithms

CNN-GCN: The CNN-GCN architecture for emotion recognition integrates the analysis of facial appearance with geometry by combining the representation learning power of convolutional neural networks and graph convolutional networks. While graph-based representation learning improves the robustness of the neuro-dynamic facial coding, the latter not only enhances the feature consistency but also improves the classification reliability under various facial expression settings where the convolutional learning extracts discriminative face features.

1) *VGG16:* VGG16 is an effective model that learns the emotional appearance patterns of facial images by applying the deep convolutional process to extract the hierarchical facial features. With the incrementally visual representation learning, the deep convolutional process and the organized architecture of VGG16 improves both the classification stability and generalization ability and ensures the reliable performance of the emotion recognition.
2) *VGG19:* By employing deeper convolutional representation learning to capture intricate emotional changes and high-level face texture information, VGG19 improves facial emotion recognition. More architectural depth enhances the capacity to extract discriminative features, enabling more thorough emotional analysis and better classification consistency in recognition tasks.
3) *ResNet18:* ResNet18 in facial emotion recognition utilizes residual learning approach to strengthen the feature propagation and deep representation learning. To ensure high accuracy classification and computational efficiency of emotional face feature extraction, residual connections can mitigate the performance degradation in the acoustic features.

4) *MobileNet:* In the context of resource constrained and real-time applications, we propose a computationally efficient and light weight face emotion identification framework, called MobileNet. Through the optimization of the convolution operations, the overall processing complexity is reduced and faster emotion predictions are enabled, along with the increased efficiency during the time of live monitoring.
5) *ConvNeXt:* ConvNeXt exploits state-of-the-art spatial representation and convolutional feature learning to improve facial emotion recognition. Greater stability of category behavior across degrees of facial expression and environmental variations is achieved through superior design, resulting in stronger emotional facial pattern representation and better generalization performance.
6) *EfficientNet:* EfficientNet achieves a high emotion identification performance by balancing expanse, depth, and input resolution. It is suitable for accurate and real-time facial emotion identification applications. The performance can be more reliable by learning effective feature representation and computational overhead can be probably reduced.

E. Integration of XAI & Flask Framework

To increase the transparency and interpretability of face expression recognition, Explainable Artificial Intelligence (XAI) was employed. Crucial facial areas for emotional prediction decisions were highlighted by visual explanation methods. The improved analytical transparency and trustworthiness during emotion analysis tasks is further strengthened by interpretability support that helps users to better understand the behavior of the prediction model. Visualization outputs also provide the following advantage: the discovery of facial attributes that most affect the categorization result, which are relevant to emotional perception.

This paper describes the design, development, and application of a web application using Flask framework. By connecting the user interface services with the emotion recognition module, we succeed in building a web application platform that provides features such as authentication, image upload, webcam monitoring, and prediction display. It provides an interactive environment for uploading an image or having a live look at the camera and monitoring the user's facial expressions. The web application also enhances the user experience of developing smart emotion-aware applications, especially the ones requiring real-time monitoring and emotion prediction.

IV. EXPERIMENTAL RESULTS

1) *Accuracy:* The aggregate percentage of correctly categorized emotion samples among all the predictions made during evaluation is known as accuracy. It shows how well the emotion detection framework works overall in accurately identifying facial expressions in a variety of emotional categories.

$$Accuracy = \frac{\mathrm{TP + TN}}{\mathrm{TP + FP + TN + FN}} \qquad (1)$$

2) *Precision:* Among all samples allocated to a certain emotional category, precision assesses the percentage of accurately predicted emotion samples. Improved reliability of emotion classification performance during facial expression analysis and fewer false positive predictions are shown by higher precision.

$$Precision = \frac{\mathrm{TP}}{\mathrm{TP + FP}} \qquad (2)$$

3) *Recall:* Recall gauges the framework's ability to accurately recognize real emotion samples that fall into each emotional category. Improved sensitivity and efficacy in identifying facial expressions without overlooking pertinent emotional examples are indicated by higher recall scores.

$$Recall = \frac{\mathrm{TP}}{\mathrm{TP + FN}} \qquad (3)$$

4) *F1-Score:* The harmonic balance between precision and recall is represented by the F1-score, which offers a thorough assessment of categorization performance. When balanced evaluation of prediction reliability is needed for emotion detection tasks including class imbalance, it is very helpful.

$$F1\,Score = 2 * \frac{Precision \times Recall}{Precision + Recall} \qquad (4)$$

Table.1 Performance Evaluation

ML Model	Accuracy	Macro F1 Score	Recall	Precision
CNN-GCN	0.992	0.992	0.992	0.992
VGG16	0.980	0.980	0.980	0.982
VGG19	0.967	0.968	0.967	0.971
ResNet18	0.996	0.996	0.996	0.996
MobileNet	0.988	0.988	0.988	0.988
ConvNeXt	0.976	0.976	0.976	0.977
EfficientNet	0.996	0.996	0.996	0.996

Table.1 shows how various deep learning models for facial emotion identification perform in comparison. When compared to other studied designs, ResNet18 had greater overall performance because of its robust generalization capabilities, enhanced gradient propagation, and effective residual feature learning, which allowed for more accurate emotion classification.

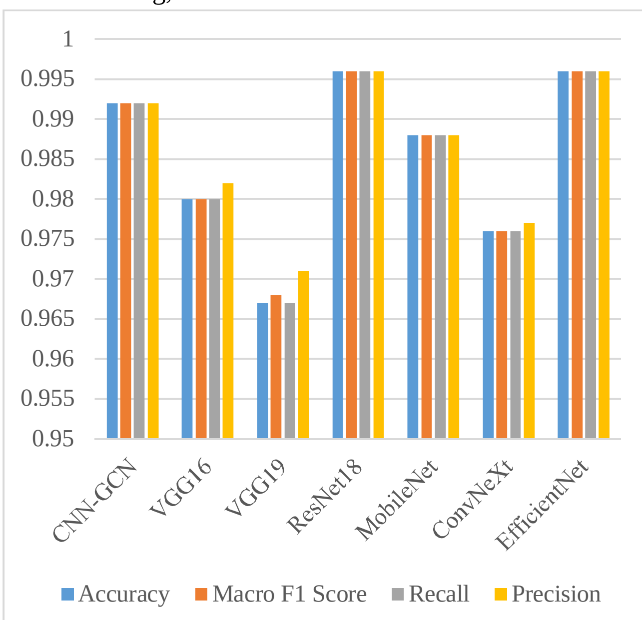


Fig.2 Comparison Graph

Fig.2 shows a graphical comparison of various emotion recognition algorithms using measures including accuracy, precision, recall, and F1-score. The animation displays the efficacy of deep feature learning techniques for attaining dependable facial emotion recognition performance and clearly highlights performance differences between models.

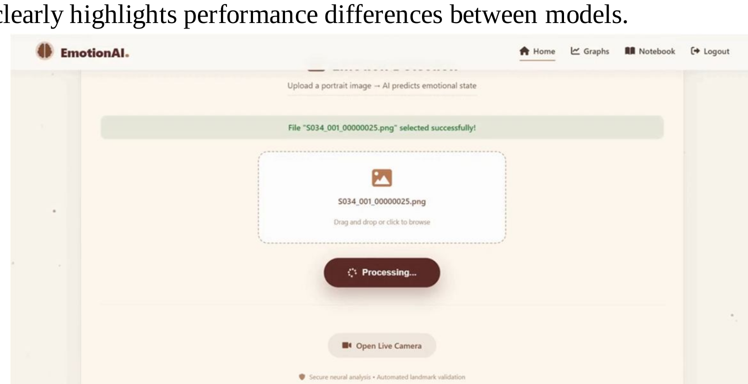


Fig.3 Image Upload Interface

Fig.3 shows the UI for uploading images created for facial expression analysis. In addition to supporting interactive access to emotion categorization, visualization outputs, and real-time analytical features within the application environment, the interface enables users to contribute facial photos for prediction.

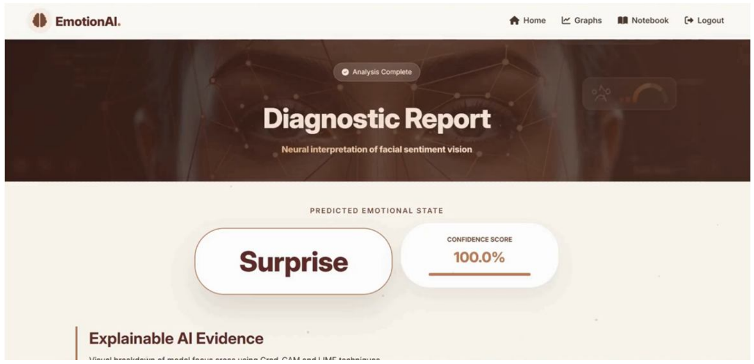


Fig.4 Emotion Prediction Result for Surprise Expression

Fig.4 displays the emotion detection results produced for a picture of a surprised face. The framework exhibits accurate facial expression analysis and dependable prediction capacity by effectively identifying the emotional category and displaying the related confidence score.

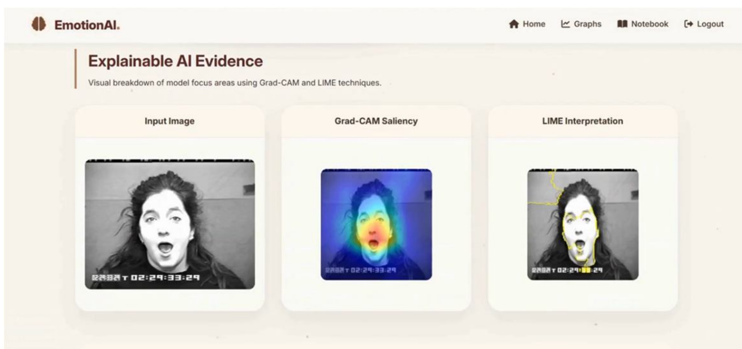


Fig.5 Explainable AI Visualization for Surprise Emotion

Fig.5 shows the explainable AI graphic produced for unexpected emotion prediction. Prediction analysis improves interpretability, transparency, and analytical comprehension of the emotion recognition behavior by highlighting key face regions that influence categorization judgments.

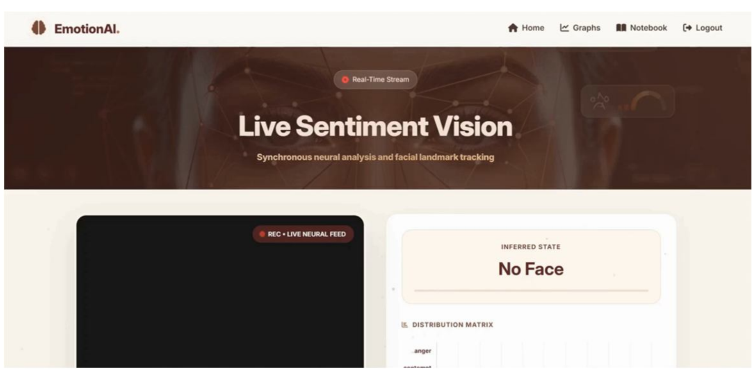


Fig.6 Real-Time Emotion Recognition Interface

Fig.6 shows how live webcam-based facial analysis is supported with a real-time emotion recognition interface. Stable emotion detection performance is made possible by continuous facial monitoring and instantaneous prediction production, which enhances the usefulness of interactive and intelligent emotion-aware application settings.

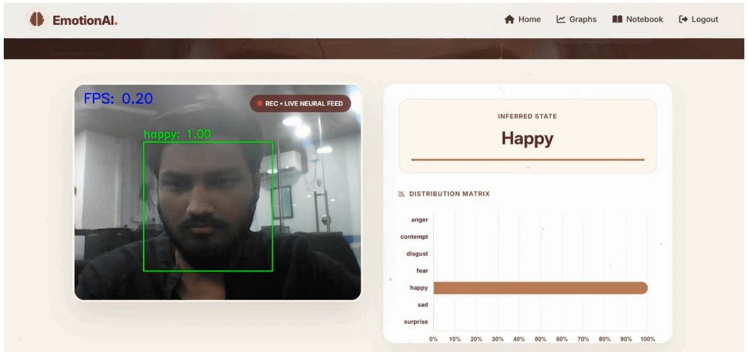


Fig.7 Emotion Prediction Result for Happy Expression

Fig.7 displays the emotion recognition result for a joyful look on the face. The framework demonstrates dependable facial expression understanding and efficient real-time emotion analysis capacity by correctly classifying the observed emotion and producing confidence-based prediction output.

V. CONCLUSION

Facial emotion recognition plays a crucial role in enabling intelligent human–computer interaction, healthcare support, behavioral analysis, and automated monitoring applications. In this work, a real-time emotion recognition framework was developed by combining a pretrained ResNet-based Convolutional Neural Network (CNN) with a Graph Convolutional Network (GCN) to capture both facial appearance features and the structural relationships among 468 facial landmarks extracted using MediaPipe Face Mesh. The model was trained and evaluated on the CK+ dataset, which consists of seven emotion categories. To improve feature representation and classification performance, a hybrid loss function incorporating cross-entropy loss and geometric loss was employed, while adversarial feature regularization enhanced the model's robustness. Explainable AI techniques, including Grad-CAM and LIME, were integrated to provide greater transparency and interpretability of the predictions. The framework was further deployed as a Flask-based web application capable of performing both real-time webcam-based and image-based emotion recognition. Experimental results demonstrated the effectiveness of the proposed approach, achieving an accuracy of 99.2% with the CNN-GCN model, while EfficientNet-B0 achieved the highest accuracy of 99.6%. Overall, the proposed system offers an accurate, reliable, and interpretable solution for real-time emotion-aware intelligent applications.

VI. FUTURE SCOPE

Real-world datasets of larger sizes with more diverse lighting conditions, postures and occlusions could be utilized to further develop the proposed emotion identification framework to provide an overall robustness. For more accurate affective computing, addition of multimodal emotion detection methods such as speech, textual and physiological signals could further push it. The integration with transformer models and the more complex graph neural networks may be utilized to enrich the temporal emotion understanding and contextual feature extraction from the video sequence. Further utilization on mobile and edge devices could broaden the accessibility for human-robot interaction, driver safety, smart classroom and healthcare analytics. The further refinement may also enhance the computational efficiency and real-time capability.

VII. ACKNOWLEDGEMENT

I express my sincere gratitude to my guide, Dr. B. Kranthi Kiran, Professor, Department of Computer Science and Engineering, Jawaharlal Nehru Technological University Hyderabad, for his valuable guidance, continuous support, and constructive suggestions throughout the development of this research work.

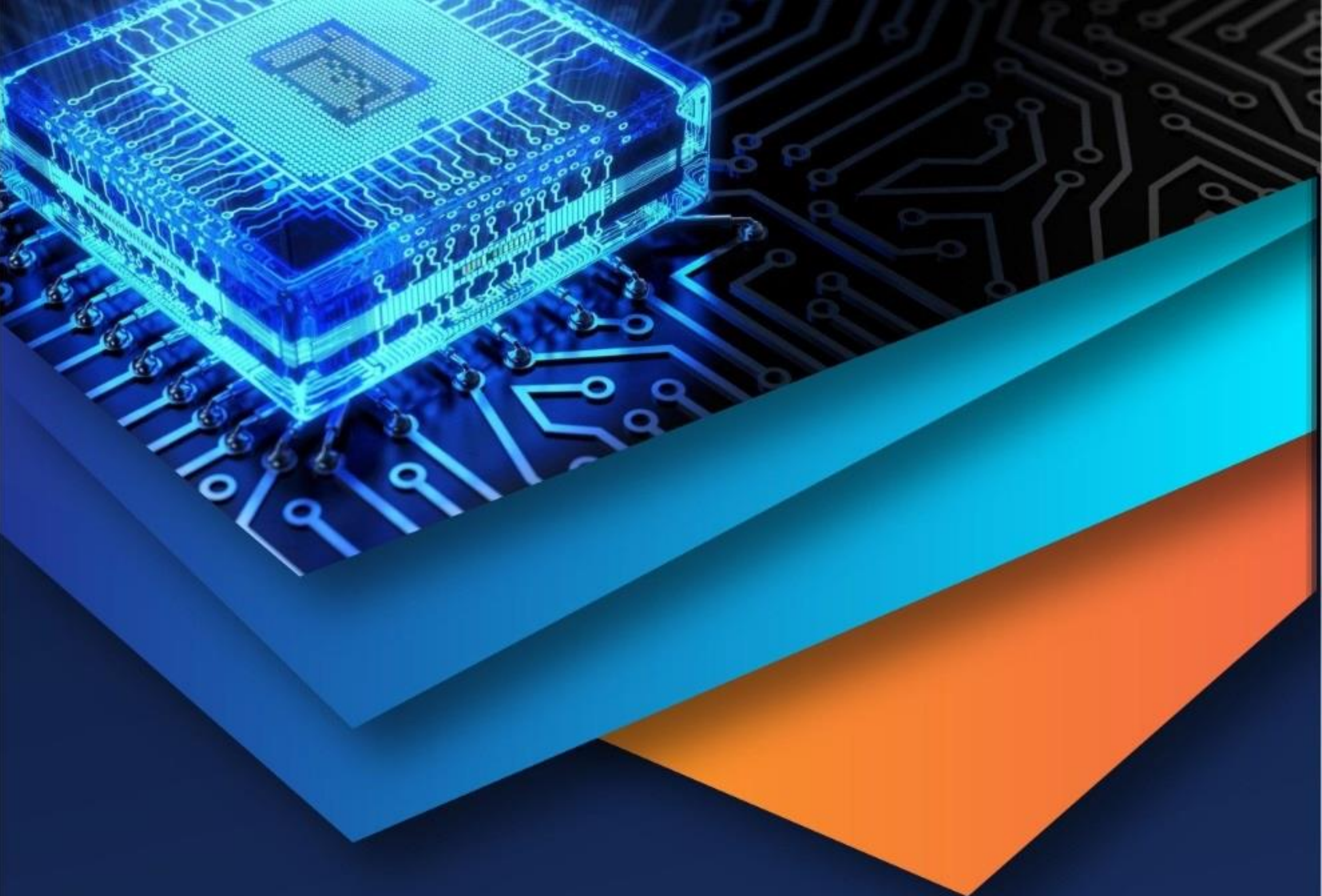
REFERENCES

[1] J. Xu, H. Si, H. Wang, S. Li, H. Wang, W. Song, and R. Zhu, “ELG: Emotion Recognition Convolutional Model Integrating Local and Global Facial Features,” in 2024 IEEE International Conference on High Performance Computing and Communications (HPCC), 2024, pp. 61–68.

[2] Y. Zhang, J. Li, and M. Noto, “Application of a Hybrid CNN-SE-GCN Model with Multi-Loss Optimization and Adversarial Samples in Real-Time Emotion Recognition,” in Proceedings of the 2025 14th International Conference on Software and Computer Applications, 2025, pp. 191–197.

[3] D. Li, J. Sun, W. Liu, L. Wang, and N. Zhou, “Dense-Attention CNN with Spatial-Attention Fusion for Robust Facial Expression Recognition,” Information Technology and Control, vol. 54, no. 3, pp. 844–863, 2025.

[4] M. Karnati, A. Seal, D. Bhattacharjee, A. Yazidi, and O. Krejcar, "Understanding Deep Learning Techniques for Recognition of Human Emotions Using Facial Expressions: A Comprehensive Survey," IEEE Transactions on Instrumentation and Measurement, vol. 72, pp. 1–31, 2023.
[5] X. Jin, Z. Lai, and Z. Jin, "Learning Dynamic Relationships for Facial Expression Recognition Based on Graph Convolutional Network," IEEE Transactions on Image Processing, vol. 30, pp. 7143–7155, 2021.
[6] F. Zhang, T. Zhang, Q. Mao, and C. Xu, "Geometry Guided Pose-Invariant Facial Expression Recognition," IEEE Transactions on Image Processing, vol. 29, pp. 4445–4460, 2020.
[7] F. M. A. Mazen, A. A. Nashat, and R. A. A. A. Seoud, "Real-Time Face Expression Recognition Along With Balanced FER2013 Dataset Using CycleGAN," International Journal of Advanced Computer Science and Applications, vol. 12, no. 6, 2021.
[8] M. Aly and N. S. Alotaibi, "A Comprehensive Deep Learning Framework for Real-Time Emotion Detection in Online Learning Using Hybrid Models," Scientific Reports, 2025.
[9] M. A. Sarikaya and G. Ince, "Improved BCI Calibration in Multimodal Emotion Recognition Using Heterogeneous Adversarial Transfer Learning," PeerJ Computer Science, vol. 11, p. e2649, 2025.
[10] H. A. Shehu, W. Browne, and H. Eisenbarth, "An Adversarial Attacks Resistance-Based Approach to Emotion Recognition From Images Using Facial Landmarks," in 2020 29th IEEE International Conference on Robot and Human Interactive Communication (RO-MAN), 2020, pp. 1307–1314.
[11] H. Burrows, J. Zarrin, L. Babu-Saheer, and M. Maktab-Dar-Oghaz, "Realtime Emotional Reflective User Interface Based on Deep Convolutional Neural Networks and Generative Adversarial Networks," Electronics, vol. 11, no. 1, p. 118, 2021.
[12] V. Janjanam, D. M. Thallam, B. G. P. Spoorthi, and K. G. Suma, "Learning-Based Real-Time Emotion Analytics for Multi-Person Settings With Confidence Visualization," in 2026 IEEE International Conference on AI Engineering and Innovations (AIEI), 2026, pp. 1–6.
[13] M. Sadiq, Y. Zhang, Y. Zhou, M. Mahmud, M. Azhar, M. Durad, and J. Liang, "A Context-Aware Dropout-Based Occlusion-Adaptive Network for Robust Facial Landmark and Emotion Detection," Journal of King Saud University Computer and Information Sciences, 2026.
[14] L. R. Ropero, J. De Laet, F. Lemic, P. S. Nácher, N. N. Bhat, S. Abadal, and X. Costa-Pérez, "Towards Emotion Recognition With 3D Pointclouds Obtained From Facial Expression Images," IEEE Transactions on Affective Computing, 2026.
[15] M. Dwijayanti, M. Iqbal, and B. Y. Suprapto, "Real-Time Implementation of Face Recognition and Emotion Recognition in a Humanoid Robot Using a Convolutional Neural Network," IEEE Access, vol. 10, pp. 89876–89886, 2022.
[16] J. Shan and S. Eliyas, "Exploring AI Facial Recognition for Real-Time Emotion Detection: Assessing Student Engagement in Online Learning Environments," in 2024 3rd International Conference on Artificial Intelligence for Internet of Things (AIIoT), 2024, pp. 1–6.
[17] R. R. Selvaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh, and D. Batra, "Grad-CAM: Visual Explanations From Deep Networks via Gradient-Based Localization," in Proceedings of the IEEE International Conference on Computer Vision (ICCV), 2017, pp. 618–626.
[18] M. T. Ribeiro, S. Singh, and C. Guestrin, "Why Should I Trust You? Explaining the Predictions of Any Classifier," in Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 2016, pp. 1135–1144.
[19] C. Lugaresi, J. Tang, H. Nash, C. McClanahan, E. Uboweja, M. Hays, F. Zhang, C. Chang, M. G. Yong, J. Lee, et al., "MediaPipe: A Framework for Building Perception Pipelines," arXiv preprint arXiv:1906.08172, 2019.
[20] P. Lucey, J. F. Cohn, T. Kanade, J. Saragih, Z. Ambadar, and I. Matthews, "The Extended Cohn-Kanade Dataset (CK+): A Complete Dataset for Action Unit and Emotion-Specified Expression," in 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops, 2010, pp. 94–101.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY